

노드 당 다중 스패닝 트리를 이용한 2계층 라우팅

정회원 서창진*, 신지수**, 준회원 김경미***

Layer 2 Routing with Multi-Spanning Tree Per a Node

Changjin Suh*, Jisoo Shin** *Regular Members*, Kyungmi Kim*** *Associate Member*

요약

캐리어 이더넷은 여러 개의 2계층 매트릭스 망을 연결하는 광대역 백본망이다. 백본망에서는 복수 개의 스패닝 트리를 구성하는 다중 스패닝 트리 프로토콜(Multiple Spanning Tree Protocol : MSTP)을 사용하는데 최근에 간단하면서도 성능이 더 우수한 최단 경로 브리징(IEEE802.1aq - Shortest Path Bridging : SPB)이 제안되었다. SPB는 목적지 노드 당 한 개씩 생성한 최단 경로 스패닝 트리를 따라서 프레임을 전송한다. 그러나 불행하게도 SPB는 목적지 노드마다 한 개의 경로만 설정하기 때문에 망의 상태 변화에 대처하기 위해서 라우팅 경로를 변경할 수 없다. 만일 목적지 노드별로 복수 개의 경로를 설정한다면 망의 상태에 적응하는 라우팅 방식을 구현할 수 있다. 본 논문은 이 철학을 적용한 라우팅 방식인 목적지 에지 노드 분할을 이용한 스패닝 트리(Edge Node Divided Spanning Tree : ENDIST)를 제안한다.

ENDIST는 에지 노드를 연결된 링크의 수만큼의 서브 노드로 분할하고 각 서브 노드마다 SPB를 사용하여 스패닝 트리를 생성한 후에, 목적지 노드로 향하는 복수 개의 라우팅 경로 중에서 한 개의 라우팅 경로를 망이나 노드의 상태에 따라 플로우 단위로 선택한다. 추가된 트래픽 엔지니어링 기법으로 우리는 백본망에서 지연 시간을 낮추고 처리율(throughput)을 향상시킬 수 있다. 시뮬레이션을 통해 우리는 최대 부하에서 ENDIST의 처리율이 STP(Spanning Tree Protocol)보다 약 3.4-5.8배, SPB보다 약 1.5-2.0배로 우수하게 동작함을 확인하였다. 특히, ENDIST의 처리율은 우리가 조사한 경우 중 절반에서 별도의 논문에서 이론적으로 계산한 한계치에 매우 근접함을 확인하였다. 이는 ENDIST가 2계층 라우팅의 성능을 획기적으로 향상시켜서 거의 완벽한 스패닝 트리 계열의 라우팅 방식이 되었음을 뜻한다.

Key Words : Carrier Ethernet, STP, MSTP, SPB, MMSTP

ABSTRACT

Carrier Ethernet backbone network integrates distributed layer-2 based metro networks. In this networks, Multiple Spanning Tree Protocol (MSTP) has been used as a main routing protocol that allows multiple spanning trees in a network. A better routing protocol called IEEE802.1aq - Shortest Path Bridging (SPB) is recently proposed, that generates the shortest spanning tree per a destination node. As SPB provides a routing path per a destination node, there is no way to adapt network traffic at normal condition. If we are free from the principle of "a spanning tree per a destination node", we can achieve adaptive routing. Based on this philosophy, we propose a new spanning tree based protocol - Edge Node Divided Spanning Tree (ENDIST).

※ 본 연구는 한국전자통신연구원 위탁연구과제(5010-2008-0086) 지원 및 숭실대학교 교내연구비 지원으로 이루어졌음
* 숭실대학교 컴퓨터학부 교수(cjsuh@ssu.ac.kr), ** 숭실대학교 대학원 컴퓨터학과 박사과정(jsshin@networking.com)
*** 숭실대학교 대학원 컴퓨터학과 석사(elmo59@gmail.com)

논문번호 : KICS2008-05-208, 접수일자 : 2008년 5월 13일 최종논문접수일자 : 2008년 8월 18일

ENDIST divides an edge node into sub-nodes as many as connecting links from the node and each sub-node generates a single shortest path tree based on SPB. Depending on network or nodal status, ENDIST chooses a better routing path by flow-basis. This added traffic engineering ability contributes to enhanced throughput and reduced delay in backbone networks. The simulation informs us that ENDIST's throughput under heavy load performs about 3.4-5.8 and 1.5-2.0 times compared with STP's and SPB's one respectively. Also, we verified that ENDIST's throughput corresponds to the theoretical upper bound at half of cases we investigated. This means that the proposed ENDIST is a dramatically enhanced and the close-to-perfect spanning tree based routing schemes.

I. 서 론

1970년대 초에 개발된 이더넷은 실제적인 표준을 넘어서 현재 범용 랜의 유일한 표준으로 사용되고 있다. 1990년 이후 이더넷의 속도는 100Mbps에서 1Gbps, 10Gbps로 향상되었다. 속도 면에서 뿐만 아니라 이더넷은 사용범위에 있어서도 2000년대에 매트릭스 이더넷을 성공시켰으며 이제는 그 영역을 백본망으로 넓히고 있다.

사업자 백본 브리지 망(Provider Backbone Bridged Network : PBBN)은 IEEE802.1ah에서 정의한 캐리어 이더넷 브리지를 사용하여 물리적으로 떨어진 2계층 브리지 망(매트릭스)을 연결한다^[1,2]. PBBN은 백본 에지 브리지(Backbone Edge Bridge : BEB)와 백본 코어 브리지(Backbone Core Bridge : BCB)로 구성된다. BEB는 망의 에지에 위치하여 라우팅과 프레임 형식 변환을 담당하며, BCB는 망의 코어에서 대규모로 프레임을 전송한다. BEB는 여러 개의 2계층 브리지 망을 연결하며, 코어 망은 여러 BEB들을 연결시킨다. BEB가 수신한 프레임은 PBBN을 통해서 다른 BEB로 전달된다. BEB 중에서 상위 망을 연결하는 브리지인 게이트웨이 브리지는 예외적으로 코어 망에 위치한다. 이후에서는 BEB는 에지 노드로, BCB는 코어 노드로, PBBN은 백본망으로 각각 명명한다. 또한 게이트웨이 브리지를 제외한 모든 에지 노드를 순수 에지 노드라고 부른다.

본문에 앞서 우리는 2계층 라우팅 프로토콜을 3세대로 구분하여 설명한다. 1세대인 STP는 망 전체에서 단 한 개의 스패닝 트리를 사용한다. 실제 동작할 때에는 스패닝 트리에 속하는 능동(active) 링크만이 사용되어서 비능동 링크는 데이터 전달용에서 제외된다. 랜에서 사용하는 이러한 간단한 방식은 값비싼 링크가 포설된 백본망에는 적합하지 않다. 2세대 라우팅인 MSTP는 설치된 모든 링크를 사용하기 위해서 한 개의 망에서 여러 개의 STP를

사용한다. MSTP에서 하나의 링크는 여러 개의 STP에 대해서 능동 링크로 동작하므로 모든 링크는 골고루 사용되어서 처리율(throughput)을 향상시킨다. 각 목적지 노드마다 하나의 최단 경로 스패닝 트리를 사용하는 SPB도 2세대 라우팅에 속한다. 마지막으로 3세대 라우팅은 MMSTP (Multiple MSTP)로 명명된다. 이름에서 첫 번째 M은 MSTP와 달리 한 목적지 노드에 대해서 복수 개의 STP를 부여할 수 있기 때문이다.

3세대 라우팅에는 현재로서는 이 논문이 소개할 ENDIST가 유일하다. ENDIST는 백본 토폴로지의 성질을 활용한다. 기존의 STP 기반 라우팅은 임의의 토폴로지를 가지는 랜에서 동작하도록 설계되었다. 반면에 백본망에서 노드들은 계층적인 구조를 가지고 매우 규칙적으로 배열되어 있으며 안정성을 높이기 위해서 복수 개의 링크로 연결된다는 성질을 가지고 있다. 우리는 백본망의 규칙적인 배열 때문에 주어진 출발 노드와 도착 노드에 대해서 많은 최단경로가 존재한다는 사실을 알았다. ENDIST의 핵심 아이디어는 “백본망에서 상존하는 복수 개의 최단경로를 최대 활용하자”이다.

그림 1은 소규모의 백본망의 예이다. 그림 1에서 원형 노드는 코어 노드를 나타내고 사각형 노드는 에지 노드를 나타낸다. 에지 노드 U 는 상위 레벨의 망을 연결하는 게이트웨이 노드이다. 그림 1에서 모든 순수 에지 노드인 A 에서 F 는 두 개의 링크로 코어 망과 연결된다. 그림 1의 백본망에 SPB를 적용하면 총 7 개의 에지 노드에서 7 개의 STP를 구성하며 ENDIST에서는 6 개의 순수 에지 노드와 게이트웨이 노드 U 에서 각각 2 개씩과 4 개의 STP를 구성하여 총 16 개의 STP를 생성한다. 부가적으로 ENDIST에서는 소스 에지 노드의 두 서브 노드를 연결하는 여러 개의 경로 중에서 하나의 경로를 선택하는 규칙이 필요하다. 본문에서 우리는 송수신 노드 사이의 홑 수와 소스 노드의 출력 큐의 길이에 따라 하나의 경로를 선택하는 규칙을 소개한다.

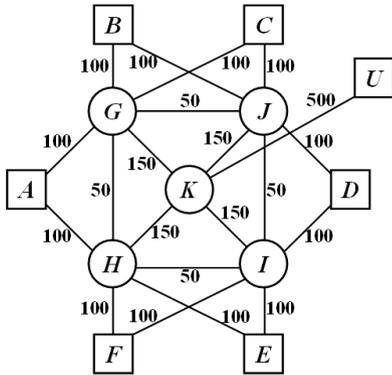


그림 1. 소규모 백본망

본 논문은 다음과 같이 구성된다. II장에서는 기존 2계층 라우팅 알고리즘을 소개하고 III장에서는 ENDIST의 경로 설정 방법과 프레임 전달하는 방법을 소개한다. IV장에서는 시뮬레이션을 통하여 ENDIST와 관련된 표준 프로토콜을 비교분석하고 V장에서 결론을 맺는다.

II. 2계층 라우팅 알고리즘

이 장에서는 2계층 라우팅 알고리즘을 3 세대로 구분하여 설명하고 논문에서 발표된 비국제표준의 스페닝 트리 라우팅 방식에 대해 살펴본다.

1세대 라우팅 방식은 스페닝 트리 프로토콜 (Spanning Tree Protocol : STP)이다³⁾. STP는 여러 개의 이더넷을 마치 하나의 이더넷처럼 동작시킨다. STP는 경로를 재구성할 때 약 30초의 시간이 필요하며 이 동안에는 망에서 데이터를 전달할 수 없다. 빠른 스페닝 트리 프로토콜(Rapid Spanning Tree Protocol : RSTP)은 STP의 재구성 시간을 100 밀리초 이하로 개선하였으며 현재에는 STP를 대체하는 단일 표준이 되었다.

2세대 라우팅 방식은 다중 스페닝 트리 프로토콜 (Multiple Spanning Tree Protocol : MSTP)로 시작되었다. MSTP는 망을 여러 개의 독립적인 지역으로 나누고 하나의 지역에 여러 개의 스페닝 트리를 생성하기 때문에 STP의 확장형으로 간주할 수 있다⁴⁾. STP와 MSTP는 필요한 에지 노드만 연결하는 부분 스페닝 트리를 정의하여 가상 랜(Virtual LAN : VLAN)⁵⁾을 생성할 수 있는데, 이를 이용하면 트래픽을 분산시키고 링크 사용률을 높일 수 있다. 그러나 MSTP는 구성 방법이 복잡하고 망의 모든 노드가 VLAN에 대한 정보를 유지해야 하기 때문에

관리가 힘들다⁶⁾. 이러한 단점을 보완하기 위해 최단 경로 브리징 기술(Shortest Path Bridging : SPB)⁶⁾이 최근에 제안되었다. SPB는 에지 노드마다 하나의 스페닝 트리를 생성하는 기법으로 동작은 매우 간단하면서도 전송 효율이 매우 높다.

3세대 라우팅으로 분류한 ENDIST는 에지 노드를 연결된 링크 개수만큼의 서브 노드로 분할하고 각 서브 노드를 루트로 하는 복수 개의 SPB를 생성한다. 3세대 라우팅은 주어진 목적지에 대해서 망이나 노드의 상태에 따라서 선택가능한 여러 라우팅 경로 중에 한 경로를 선택하기 때문에 부하를 균등하게 분산시킬 수 있다.

SmartBridge⁷⁾는 망의 세그먼트와 브리지의 연결 형태에 따라 상이하게 동작하는 브리지 구조를 제안하였다. SmartBridge에서 모든 프레임은 토폴로지 습득 기법을 통해 최단 경로를 따라 전송된다. SmartBridge는 낮은 지연의 경로와 짧은 재구성 시간을 제공한다.

STAR(Spanning Tree Alternate Routing)⁸⁾는 기본적으로 스페닝 트리를 사용하며 식별한 대체 경로를 사용할 수 있을 경우 대체 경로로 프레임을 전달한다. STAR는 송수신 노드 사이의 지연을 감소시키지만 한 경로 내에서 일어나는 과부하 문제는 해결하지 못한다.

Viking⁹⁾은 VLAN을 이용한 다중 스페닝 트리를 만들어서 복수 개의 링크를 제공하여 망의 성능을 향상시키고 백업 경로를 미리 설정하여 재구성 시간을 줄인다. 그러나 Viking은 백업 경로 생성을 위한 별도의 서버가 필요하다.

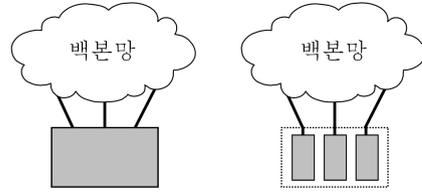
이외에도 스페닝 트리의 QoS를 높이기 위해 MSTP를 변형한 프로토콜¹⁰⁾과 트래픽 엔지니어링을 지원하기 위해 STAR에서 프레임 우선순위를 사용하는 방법¹¹⁾, 링크의 가중치를 이용하여 스페닝 트리를 구성하는 연구¹²⁾ 등이 제안되었다. 이들 모두는 국제 표준인 SPB와 비슷한 성능을 보이는 정도이며 2계층 라우팅 방식이라는 한계 때문에 본 논문에서 소개할 ENDIST에 비해서는 성능이 매우 뒤떨어진다.

III. ENDIST

이 장에서는 ENDIST를 정의한다. 3.1절에서는 에지 노드의 구조를 정의하고 ENDIST의 동작을 설명한다. 3.2절에서는 다중 경로를 생성하고 최종 경로를 선택하는 방법을 정의한 후 3.3절에서 예를 들어서 설명한다.

3.1 에지 노드의 정의 및 구현

ENDIST에서 에지 노드는 자신과 연결된 링크 개수만큼의 서브 노드로 분할된다. 그림 2는 세 개의 링크로 백본 코어 망과 연결된 에지 노드가 세 개의 서브 노드로 분할된 모습을 보여준다. 그림 2.(b)에서 보듯이 각각의 서브 노드는 링크와 일대일로 연결되며 다른 링크와는 분리된다. 그 결과 특정한 서브 노드를 선택하면 첫째 홉(hop)으로 사용할 링크가 자동적으로 결정된다.



(a) 3개의 링크로 연결된 에지 노드 (b) 3개의 서브 노드로 분할된 에지 노드

그림 2. 에지 노드의 분할

ENDIST가 동작하기 위해서 우리는 서브 노드 단위의 주소가 필요하다. 예를 들어 세 개의 서브 노드를 가진 에지 노드는 MAC 주소의 가장 마지막 두 개의 비트를 제외하고 동일한 MAC 주소를 가지며, 마지막 두 개의 비트는 서브 노드의 식별자로 사용된다. 백본망에서는 낮은 레벨의 망으로부터 수신된 프레임에 새로운 송수신 MAC 주소를 추가하는 MAC-in-MAC 기법을 사용한다. 이처럼 백본망에서는 새로운 주소가 사용되는 과정에서 주소 필드와 서브 노드 필드를 구분하여 사용할 수 있다.

3.2 사용할 경로와 서브 주소의 선택

ENDIST는 하나의 목적지 노드에 대해서 복수 개의 경로를 설정할 수 있고 라우팅 경로는 플로우 단위로 결정된다. 송신 에지 노드에서 새로운 플로우가 발생하면 사용할 스페닝 트리를 선택하고 플로우가 모두 전송될 때까지 고정된 경로를 사용한다. 영속적으로 사용하는(permanent) 플로우는 관리자의 도움으로 다음과 같이 스페닝 트리를 변경할 수 있다. 관리자는 경로를 바꾸고자 하는 플로우에 새로운 경로를 마련한 후 새로운 플로우 번호를 부여한다. 또한 기존 플로우와 경로의 사용을 중지하고 새로운 경로를 따라서 새로운 플로우 번호로 전송한다. 플로우 교체에 앞서서 관리자는 모든 에지 노드에 플로우 번호 교체를 통보하여 모든 수신 노드는 새로운 플로우를 기존 플로우와 동일하게 취급하게 한다.

이제부터 ENDIST에서 스페닝 트리를 설정하는 방법을 설명한다. ENDIST는 SPB를 이용하여 서브 노드에서 모든 목적지 서브 노드로 향하는 경로를 생성한다. 만일 출발지와 도착지 노드에서의 서브 노드 개수가 각각 $s(S)$ 와 $s(D)$ 일 때에 서브 노드 단위로 스페닝 트리를 생성하면 송신 노드에서 수신 노드로 전달할 수 있는 경로는 총 $s(S) \cdot s(D)$ 개이며 이 중에서 한 경로를 선택한다. 최종 경로의 선택은 백본망의 정책에 따라 결정될 수 있는데 ENDIST

는 간단한 두 가지 정책으로 최종 경로를 선택한다. ENDIST는 아래의 단계 1-3에 따라서 사용할 복수 개의 경로를 설정한 후 단계 4에서는 복수 개의 경로 중에서 사용할 경로를 플로우 단위로 선택한다.

- 단계 1 : 모든 에지 노드를 서브 노드로 분할하고 서브 노드 단위로 출발/도착 노드를 정하고 SPB를 이용하여 스페닝 트리를 생성한다.
- 단계 2 : 주어진 서브 노드에서 상대방 에지 노드로 향하는 $s(D)$ 개의 경로 중 최단 경로를 한 개씩 선택한다.
- 단계 3 : 단계 2에서 얻어진 경로 중에서 사용할 스페닝 트리를 선택한다. 송신 노드는 각 수신 에지 노드로 향하는 최단 홉 경로를 선택한다.
- 단계 4 : 최단 홉이 동일할 경우에는 송신 에지 노드의 출력 큐의 길이가 가장 짧은 경로를 선택한다. 이때 출력 큐는 모든 목적지 주소에 대해서 단일 큐를 사용한다.

출발노드 S 에서 도착노드 D 까지의 경로 개수는 단계 1에서 총 $s(S) \cdot s(D)$ 개가 제공된다. 단계 2에서는 이들을 출발 서브 당 모든 목적 노드를 연결하는 스페닝 트리를 만든다. 그 결과 목적지 노드 당 $s(S)$ 개의 경로가 존재한다. 단계 2에서 얻은 최단 경로는 첫 번째 사용 링크가 고정된 상태에서의 최단 경로이기 때문에 출발/도착 노드의 관점에서 보면 경로 길이가 상이하하다. 단계 3에서는 이러한 점을 고려해서 최단경로를 제공하는 스페닝 트리를 선택한다.

본 논문에서는 가장 간단한 구현 방법을 사용하기 위해서 단계 3에서는 최단 홉을, 단계 4에서는 최단 출력 큐를 경로 선정의 기준으로 삼았다. 이 방법은 시뮬레이션 결과에서 보이듯이 우수한 방법이지만 망 관리자가 자신에게 더 알맞은 정책을 사용하길 원한다면 단계 3과 4를 수정하여 개선할 수

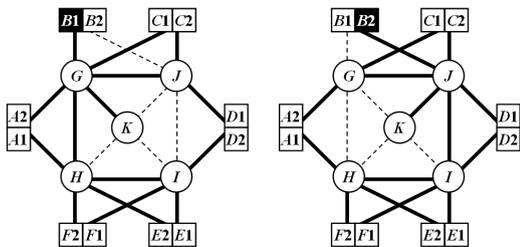
있다. 다만 단계 3에서 모든 순수 에지 노드는 목 적지 순수 에지 노드 당 평균적으로 1.5 개 이상의 충분한 다중 경로를 설정해야 본 논문에서 제시한 바와 같은 성능 향상을 이룰 수 있다.

3.3 최종 경로 선택의 예

이 절에서는 스페닝 트리를 생성하고 최종 경로를 선택하는 일례를 보여준다. 그림 3은 서브 노드 단위로 구성된 스페닝 트리를 보여준다. 그림 3.(a)는 서브 노드 B1이 루트 노드인 경우이며 그림 3.(b)는 B2가 루트 노드인 경우를 표시하고 있다. 그림 3에서 굵은 실선은 스페닝 트리에 포함된 링크이고 점선은 포함되지 않은 링크를 나타내며 루트 노드는 역상으로 표시했다. 게이트웨이 노드 U는 편의상 생략하였다. 그림 4에도 동일한 표현 방법을 사용하였다.

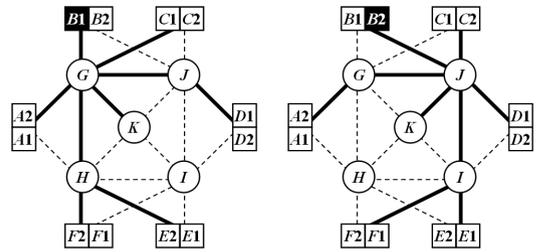
그림 3은 노드 B를 루트로 한 서브 노드 단위의 스페닝 트리이다. 그림 3.(a)의 트리는 서브 노드 B1에서 모든 서브 노드를 최단 경로로 연결한다. B1에서 C로 가는 최단 경로는 B1-G-C1과 B1-G-J-C2가 있다. B1을 사용할 경우에 길이가 긴 경로는 오히려 성능을 낮추므로 B1-G-C1만을 선택 하는 것이 바람직하다. 그 결과 그림 4.(a)에서처럼 길이가 긴 경로 B1-G-J-C2는 생략된다. 다른 에지 노드에 대해서도 동일한 방법으로 경로를 삭제하면 그림 4.(a)를 얻을 수 있다.

단계 2를 종료하면 노드 B에서 모든 에지 노드 에 도달하는 경로가 그림 4.(a)와 (b)에 하나씩 총 두 개의 경로가 준비된다. 노드 B에서 C로 전송할 때 그림 4.(a)의 B1-G-C1과 그림 4.(b)의 B2-J-C2의 두 경로는 홉 수가 동일하므로 모두 사용할 수 있다. 반면에 노드 B에서 A로 전송할 경우에는 그림 4.(a)의 B1-G-A2와 그림 4.(b)의 B2-J-G-A2의 두 경로 중에서 홉 수가 적은 B1-G-A2만을 경로로



(a) 링크 B1-G를 사용한 경우 (b) 링크 B2-J를 사용한 경우

그림 3. 단계 1을 종료했을 때 에지 노드 B를 루트로 하는 스페닝 트리



(a) 링크 B1-G를 사용한 경우 (b) 링크 B2-J를 사용한 경우

그림 4. 단계 2를 종료했을 때 에지 노드 B를 루트로 하는 스페닝 트리

정한다. 단계 3이 적용된 결과 노드 B는 노드 A에 대해서는 4.(a)를, 노드 D에 대해서는 4.(b)만을 사용하며, 노드 C, E, F에 대해서는 4.(a)와 4.(b)를 모두 사용한다. 복수 경로가 설정된 경우에 단계 4에서 노드 B는 자신과 연결된 링크 B1-G와 링크 B2-J로 연결되는 출력 큐의 길이를 비교하여 짧은 큐 길이의 링크를 선정하며 이를 포함하는 경로를 선택하여 사용한다.

IV. 시뮬레이션 및 분석

본 장에서는 여러 개의 네트워크에서 우리가 제안한 ENDIST를 하나의 스페닝 트리를 사용하는 STP와 MSTP에 기초한 SPB, 이론적인 처리율을 계산한 이론적 한계치와 비교 분석한다. 이론적인 한계치는 maximum-flow를 LP를 이용하여 계산한 별도의 논문^[13]에서 추출했으며 나머지 세 방식의 라우팅은 NS-2^[14]를 사용하여 시뮬레이션하였다.

시뮬레이션 환경은 다음과 같다. 시뮬레이션은 그림 1과 5의 두 가지 백본망을 사용하고 모든 링크는 full-duplex의 10Gbps 링크라 가정한다. 두 그림에서 각 링크에 표시된 숫자는 링크의 전송 속도를 1Gbps 단위로 표시한 값이다. 우리는 모든 링크는 10Gbps 속도를 가진다고 가정하였으며, 가령 50은 10Gbps 링크 다섯 개가 병렬 연결되어 있다는 의미이다. ENDIST는 플로우 단위로 스페닝 트리를 선택하지만 시뮬레이션에서는 플로우 설정에 따른 영향을 없애기 위해서 프레임 단위로 스페닝 트리를 선택하였다.

시뮬레이션에서는 다음과 같이 트래픽을 가정한다. 우선 모든 순수 에지 노드는 동일한 부하의 트래픽을 생성하며 이 중에서 상향 트래픽 비율로 명명된 e에 해당하는 트래픽을 게이트웨이 노드 U로

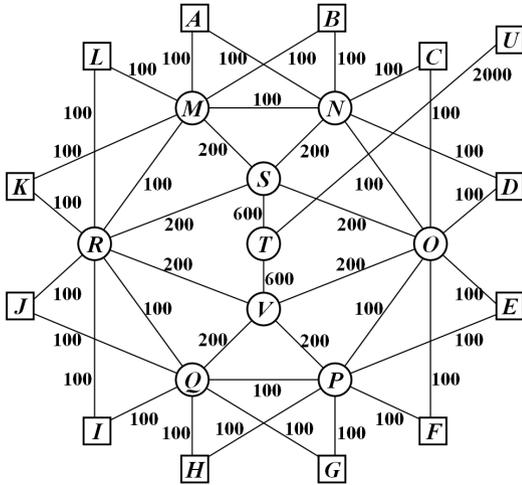


그림 5. 대규모 백본망의 예.

보낸다. 노드 U에서는 상위 망으로 출력되는 트래픽과 동일한 트래픽을 상위 망으로부터 입력한다. 나머지 (1-e)의 트래픽은 모든 인기 노드와 일반 노드에게 공평하게 분배된다. 인기 노드는 일반 노드에 비해서 두 배의 트래픽을 수신한다. 시뮬레이션에서는 인기 노드를 설정한 경우와 설정하지 않은 경우를 각각 비대칭과 대칭 트래픽 분포라고 부른다. 시뮬레이션에서 인기 노드는 전체의 순수 에지 노드의 약 1/4에 해당하는 수 - 소규모 백본망에서는 두 개, 대규모 백본망에서는 세 개 - 를 선정하였다. 트래픽이 집중되는 효과를 높이기 위해서 인기 노드는 인접한 일련의 노드로 선정하였다.

각 노드는 유한 버퍼를 가지고 있으며 버퍼 용량을 초과하는 프레임은 손실된다. 대칭적인 트래픽이 인가된 경우에 병목현상을 초래하는 링크를 통과하는 트래픽이 주로 버퍼초과로 소실되기 때문에 단대단으로 성공적으로 전달되는 트래픽은 일반적으로 대칭의 분포를 구성하지 못한다. LP로 값을 구한 이론적 한계치에서는 각 에지 노드에서 손실이 발생하지 않을 만큼의 플로우만을 발생시킨다는 가정으로 문제를 풀기 때문에 중간 노드에서의 손실은 발생하지 않는다.

위와 같은 시뮬레이션 환경에서 우리는 순수 에지 노드의 부하 발생량에 따른 처리율과 전송 지연 시간, 패킷 손실률을 계산하였다. 그림 6-8은 각각 그림 1의 소규모 백본망에서 대칭분포이고 상향 트래픽 비율 e가 20%일 때 처리율과 평균 지연 시간, 패킷 손실률을 보여준다.

그림 6은 부하의 변동에 따른 백본망의 처리율의

변화량을 보여준다. 그림 6에서 3세대 방식인 ENDIST가 이론적 한계치에 가장 근접한 우수한 성능을 보이며 1세대 방식의 STP가 가장 저조한 성능을 보여준다. 이는 사용하는 스패닝 트리의 수가 ENDIST, SPB, STP의 순서로 많기 때문이다.

그림 7과 8은 소규모 백본망에서의 상향 트래픽 비율이 20%인 대칭 부하가 인가될 때의 전송 지연과 패킷 손실률을 각각 보여준다. 그림 7에서 ENDIST는 최대 부하에서 경부하에서의 지연 시간의 두 배 이내의 지연 시간을 갖는데 이는 ENDIST가 트래픽 부하가 큰 경우에도 지연 시간이 무시할 수준임을 의미한다. STP와 SPB의 지연 시간은 중간 부하까지 급증하다가 높은 부하에서 버퍼 오버플로우가 발생하여 포화된 모습을 보인다. 만일 시뮬레이션에서 버퍼의 용량을 늘린다면

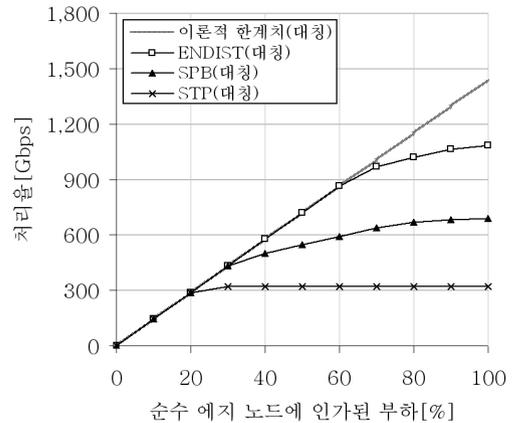


그림 6. 소규모 백본망에서 대칭, (e=0.2)인 경우의 처리율

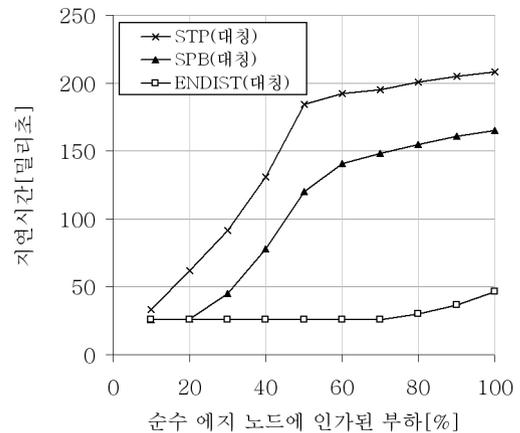


그림 7. 소규모 백본망에서 대칭, (e=0.2)인 경우의 지연 시간

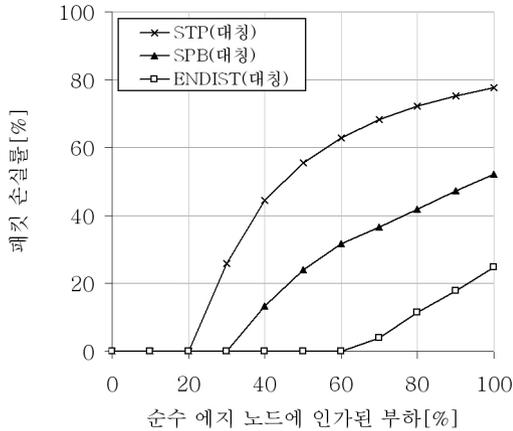


그림 8. 소규모 백본망에서 대칭, ($e=0.2$)인 경우의 패킷 손실률

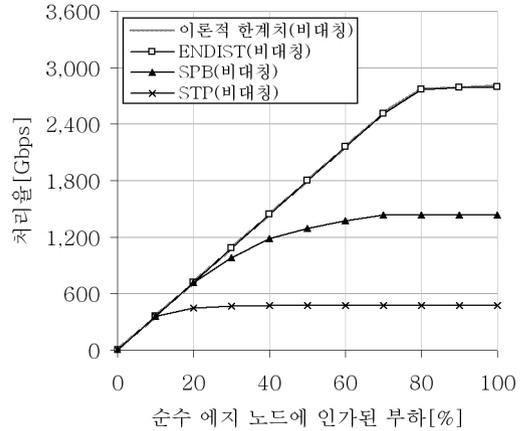


그림 10. 대규모 백본망에서 비대칭, ($e=0.5$)인 경우의 처리율

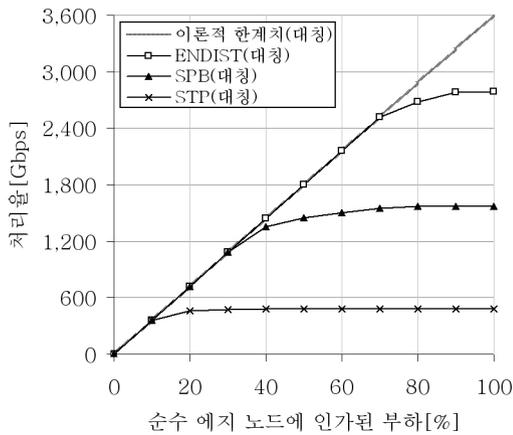


그림 9. 대규모 백본망에서 대칭, ($e=0.5$)인 경우의 처리율

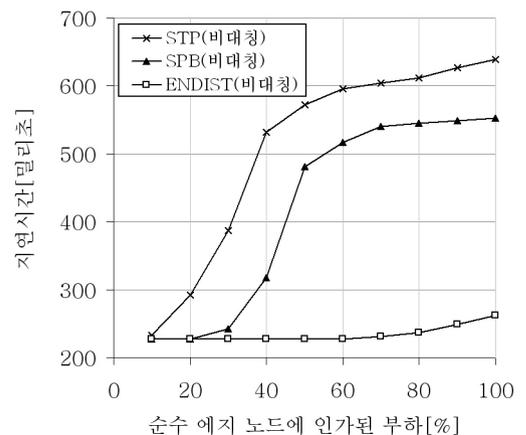


그림 11. 대규모 백본망에서 비대칭, ($e=0.5$)인 경우의 지연 시간

ENDIST와의 지연 시간의 격차는 더욱 커진다. 패킷 손실률은 처리율과 반대되는 개념이다. 그림 8에서 패킷 손실률이 나타나기 시작하는 부하의 시점은 그림 6의 처리율 그래프에서 이론적인 한계치 곡선에서 벗어나기 시작되는 시점과 동일하다. ENDIST의 패킷 손실률은 SPB나 STP보다 월등히 적다.

우리는 상향 트래픽 비율 e 가 50%일 때 대칭과 비대칭의 트래픽 분포에 대해서 실험하였다. 그림 9와 10은 대규모 백본망에서 대칭과 비대칭의 트래픽 분포에서의 처리율을 나타낸다. 그림 9에서 직선 형태를 보이던 이론적 한계치는 그림 10에서는 부분적으로 포화되었다. 그림 10에서 ENDIST는 거의 이론적 한계치와 거의 일치하고 있으므로 완벽하게 동작하고 있음을 알려준다.

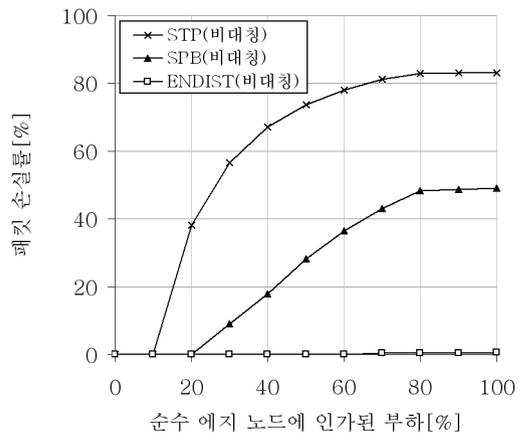


그림 12. 대규모 백본망에서 비대칭, ($e=0.5$)인 경우의 패킷 손실률

그림 11과 12는 대규모 백본망에서 e 가 50%이고 비대칭 부하일 때 발생하는 지연 시간과 패킷 손실률을 각각 보여준다. 그림 11에서 ENDIST는 사용하는 출력 큐의 길이가 짧아서 모든 부하에서 매우 낮은 지연 시간을 보여준다. 그림 12에서 STP의 패킷 손실은 부하가 10%일 때부터, SPB의 패킷 손실은 부하가 20%일 때부터 발생한다. ENDIST는 매우 미세한 패킷 손실만 발생하여 완벽에 가까운 라우팅을 수행한다.

표 1과 표 2는 실제적으로 포화 상태인 90%의 부하를 가정하고 여러 라우팅 방식의 처리율을 이론적 한계치에 대한 상대적인 값으로 보여준다. 여기에서 상황 트래픽 비율 e 를 20%와 50%로 가정하고 우리는 소규모와 대규모 백본망에서 대칭/비대칭 트래픽 분포에 대하여 조사하였다. ENDIST의 처리율은 표 1에서 SPB보다 약 1.5-2배, STP보다 약 3.4-4.3배 높으며 표 2에서는 SPB보다 약 1.6-2배, STP보다 약 3.4-5.8배 높게 나타났다. ENDIST는 소규모보다는 대규모 백본망에서, 대칭보다는 비대칭 부하에서 상대적으로 뛰어나다.

표 1. 소규모 백본망, 90% 부하에서의 처리율

e 알고리즘	20% (대칭)	20% (비대칭)	50% (대칭)	50% (비대칭)
한계치	1.00	1.00	1.00	1.00
ENDIST	0.75	0.99	1.00	1.00
SPB	0.48	0.49	0.49	0.51
STP	0.22	0.23	0.24	0.26

표 2. 대규모 백본망, 90% 부하에서의 처리율

e 알고리즘	20% (대칭)	20% (비대칭)	50% (대칭)	50% (비대칭)
한계치	1.00	1.00	1.00	1.00
ENDIST	0.54	0.75	0.78	0.99
SPB	0.33	0.45	0.44	0.51
STP	0.16	0.21	0.13	0.17

우리는 시뮬레이션을 통해 알아낸 특징을 다음과 같이 정리하였다.

- 처리율 : ENDIST는 STP보다 3배 이상, SPB보다 1.5배 이상의 성능을 나타낸다.
- 지연 시간 : ENDIST는 STP와 SPB와 비교할 때 지연 시간을 혁신적으로 줄인다.

- 패킷 손실률 : ENDIST는 STP와 SPB와 달리 극소수의 패킷만이 손실된다.
- ENDIST의 전반적인 특성 : ENDIST는 STP와 SPB보다 월등히 우수하다. 특히 ENDIST는 대칭보다는 비대칭적인 트래픽에서, 소규모보다는 대규모 백본망에서 더 우수하게 동작한다.
- ENDIST의 완벽한 정도 : ENDIST는 조사한 케이스 중 반수 정도의 경우에서 이론적인 한계치에 근접한 처리율과 패킷 손실률을 보인다.

V. 결론

우리는 기존의 스페닝 트리를 사용하는 이더넷 방식을 기존의 2 세대에서 3 세대로 확대하고 3세대 라우팅 방식의 일례를 이 논문에서 제안하였다. 이더넷의 라우팅 방식에서 1세대인 STP에서는 전체 망에서 한 개, 2세대의 SPB에서는 에지 노드 당 한 개의 스페닝 트리 사용하는데 반해서 3세대 라우팅에서는 에지 노드 당 복수 개의 스페닝 트리가 사용된다. 최초의 3세대 라우팅 방식인 ENDIST는 에지 노드를 연결된 링크 수만큼의 서브 노드로 나누고 각 서브 노드를 루트로 하는 스페닝 트리를 생성한다. 이 논문은 복수 개의 스페닝 트리의 장점을 관찰하기 위해서 매우 간단한 방식만을 적용한 상태에서 실험했지만, 본문에서처럼 STP보다 약 3.4-5.8배, SPB보다 1.5-2.0배의 처리율을 얻었고, 미미한 지연 시간만 초래하는 뛰어난 성능을 보여준다.

ENDIST는 스페닝 트리를 사용하는 간단하게 구현한 2계층 라우팅 방식으로서 이론적 라우팅의 한계치에 수렴하는 결과를 보이고 있다. 이는 3계층 라우팅에 비해서 값싸지만 성능이 낮다고 인식되는 2계층 라우팅이 이제는 3계층 라우팅 방식에 필적할 수 있음을 시사한다. 우리는 우리가 제안한 라우팅 방식의 도움으로 랜에서 시작한 이더넷이 백본망에서도 성공하여 만인이 저렴한 비용으로 인터넷을 사용하는데 일조하기를 바란다.

참고 문헌

- [1] IEEE 802.1ah/D3.0, "Virtual Bridged Local Area Networks, Amendment 6 : Provider Backbone Bridges," 2006.
- [2] IEEE 802.1aq/D0.3, "Virtual Bridged Local Area Networks, Amendment 9 : Shortest

Path Bridging,” 2006.

[3] IEEE 802.1D, “Media Access Control (MAC) Bridges,” 1998.

[4] IEEE 802.1s, “Virtual Bridged Local Area Networks, Amendment 3: Multiple Spanning Tree,” 2002.

[5] IEEE 802.1Q, “Virtual Bridged Local Area Networks,” 1998.

[6] Norman Finn, “Shortest Path Bridging-An Update on Bridging Technologies,” IEEE Tutorial, July, 2005.

[7] T. Rodeheffer, C. Thekkat, and D. Anderson, “SmartBridge: A scalable bridge architecture,” ACM SIGCOMM, 2000.

[8] K. Lui, W. Lee and K. Nahrstedt, “STAR: A transparent spanning tree bridge protocol with alternate routing,” ACM SIGCOMM, Vol.32 of 3, pp.33-46, 2002.

[9] Srikant, Kartik, Susanta and Tzi-cker, “Viking: A Multi-Spanning-Tree Ethernet Architecture for Metropolitan Area and Cluster Networks,” IEEE INFOCOM 2004.

[10] Y. Lim, H. Yu, S. Das, S. Lee, M. Gerla, “QoS-aware Multiple Spanning Tree Mechanism over a Bridged LAN Environment,” IEEE GLOBECOM 2003.

[11] K. Lui, W. Lee, K. Nahrstedt, “Link Layer Multi-Priority Frame Forwarding,” IEEE ICC 2003.

[12] M. Padmaraj, S. Nair, M. Marchetti, G. Chiruvolu, M. Ali, A. Ge, “Metro Ethernet traffic engineering based on optimal multiple spanning trees,” WOCN 2005.

[13] Changjin Suh, Shin-woong Jung, “Layer-2 Routing Analytic Model by Linear Programming,” the 10th International Conference on Advanced Communication Technology (ICACT2008) : Republic of Korea, Feb. 17~20, 2008.

[14] The ns-2 network simulator. <http://www.isi.edu/nsnam/ns>.

서 창 진 (Changjin Suh)

정회원



1982년 서울대학교 제어계측학과 졸업
 1984년 서울대학교 제어계측학과 석사
 1985년~1990년 한국전자통신연구원 선임연구원
 1996년 ECE(Electric and Computer

Engineering), Univ. of Massachusetts at Amherst 박사
 1997년~현재 숭실대학교 컴퓨터학부 교수
 <관심분야> Sensor network, Carrier Ethernet, Routing theory, BGP

신 지 수 (Jisoo Shin)

정회원



2003년 숭실대학교 컴퓨터학부 졸업
 2005년 숭실대학교 대학원 컴퓨터학과 석사
 2005년~현재 숭실대학교대학원 컴퓨터학과 박사과정

<관심분야> Sensor network, Carrier Ethernet, Routing theory 등

김 경 미 (Kyungmi Kim)

준회원



2005년 국민대학교 컴퓨터공학부 졸업
 2008년 숭실대학교 대학원 컴퓨터학과 석사
 <관심분야> Sensor network, Carrier Ethernet 등